Technology Infusion of CodeSonar into the Space Network Ground Segment (RII07)

Software Assurance Symposium

Technical Summary

Markland J. Benson, Computer Systems Manager, White Sands Complex, (575) 527-7034, Markland.Benson@nasa.gov

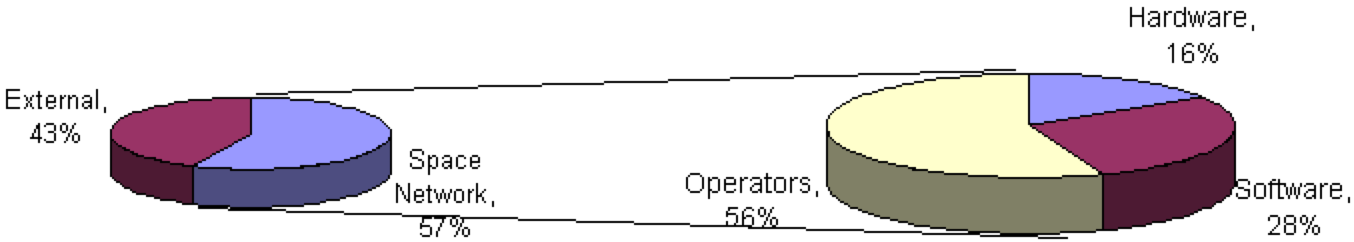
Problem Context (1)

- Ongoing maintenance by the sustaining engineering group has limited personnel and test resources but must maintain 99.9% proficiency and 97.0% availability of systems
- Major new capabilities fielded for the Space Network are often contracted to entities outside of the sustaining engineering group and historically have higher than required defect rates

$$\boxed{\frac{T_E - T_A}{T_E}}$$

Problem Context (2)

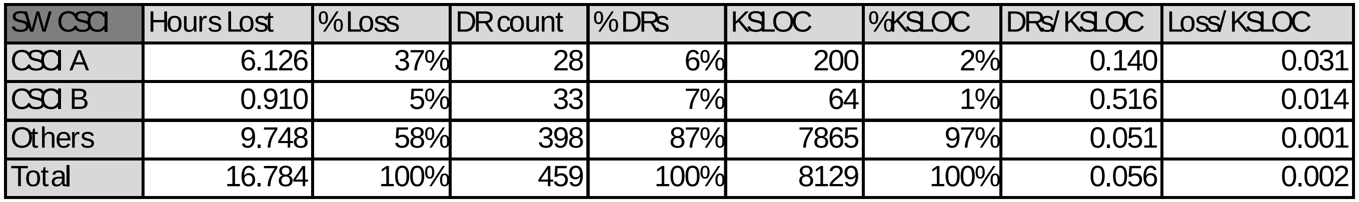
- Software has historically accounted for an annual average of 28% of the Space Network loss of availability and proficiency (low of 11% and high of 57% annually)



- CSCI A and CSCI B account for 42% of the previous eight months software data loss

Percentages reflect the portion of losses attributed to software with respect to the total loss within the control of the Space Network. For example, if the annual average loss attributable to the Space Network was 500 out of 130k total hours (99.996% proficiency), then the portion caused by software would be 500 x 0.28 = 140 hours.

Problem Context (3)



SW CSCI	Hours Lost	% Loss	DR count	% DRs	KSLOC	%KSLOC	DRs/ KSLOC	Loss/ KSLOC
CSCI A	6.126	37%	28	6%	200	2%	0.140	0.031
CSCI B	0.910	5%	33	7%	64	1%	0.516	0.014
Others	9.748	58%	398	87%	7865	97%	0.051	0.001
Total	16.784	100%	459	100%	8129	100%	0.056	0.002

- CSCI A and CSCI B have discrepancies an order of magnitude larger than other CSCIs
- CSCI A was deployed several years ago
- CSCI B was deployed eight months ago
- These data represent the past eight months

Problem Context (4)

- CSCI A represented approximately 15% of all scheduled support time over the past eight months but accounted for 37% of all loss

- CSCI B has not had schedule data long enough to trend schedule time versus loss but had delayed deployment due to problems found during system and acceptance testing

- The remainder of the Space Network systems accounted for 58% of the data loss while providing at least 75% of scheduled support

Analysis Results (1)

- Running CodeSonar and performing preliminary review of the results averaged 3.5 minutes per finding (approx. 20 hours total)
- An additional 40 hours is estimated to analyze the 37 findings deemed too complex for the initial review

- Using CodeSonar's tools to suppress known non-problems, delta tool runs will not repeat findings marked as non-problems, further reducing the time needed for review

Analysis Results (2)

- Of the 330 total findings, 70 were associated with the Standard Template Library (STL) and 109 had duplicate sources to other findings (fixing one either fixes the duplicate problems or provides a straightforward method to fix the duplicate problems)

CSCI	CSC	All	U	R	N	TP (D)	TP (P)
CSCI B	CSC A	8	0	4	1	0.500	0.625
CSCI B	CSC B	3	0	1	0	0.333	0.333
CSCI B	CSC C	12	1	2	2	0.250	0.417
CSCI B	CSC D	18	0	4	3	0.222	0.389
CSCI B	CSC E	38	0	8	0	0.211	0.211
CSCI B	CSC F	145	3	14	29	0.117	0.317
CSCI B	CSC G	2	0	0	0	0.000	0.000
CSCI B	CSC H	14	0	5	0	0.357	0.357
CSCI B	CSC I	5	0	1	0	0.200	0.200
CSCI B	CSC J	10	0	3	0	0.300	0.300
CSCI B	CSC K	46	0	2	0	0.043	0.043
CSCI A	CSC A	29	1	7	2	0.276	0.345
Total		330	5	51	37	0.170	0.282

Analysis Results (3)

CSCI	CSC	All	U	R	N	TP (D)	TP (P)
CSCI B	CSC A	8	0	4	1	0.500	0.625
CSCI B	CSC B	3	0	1	0	0.333	0.333
CSCI B	CSC C	12	1	2	2	0.250	0.417
CSCI B	CSC D	18	0	4	3	0.222	0.389
CSCI B	CSC E	38	0	8	0	0.211	0.211
CSCI B	CSC F	145	3	14	29	0.117	0.317
CSCI B	CSC G	2	0	0	0	0.000	0.000
CSCI B	CSC H	14	0	5	0	0.357	0.357
CSCI B	CSC I	5	0	1	0	0.200	0.200
CSCI B	CSC J	10	0	3	0	0.300	0.300
CSCI B	CSC K	46	0	2	0	0.043	0.043
CSCI A	CSC A	29	1	7	2	0.276	0.345
Total		330	5	51	37	0.170	0.282

- Of the remaining 151 findings, 29 are findings that could potentially be problems but the coding style and standards used prevent the problem
- An addition 29 are definite false positives, which were determined to not be problems upon review

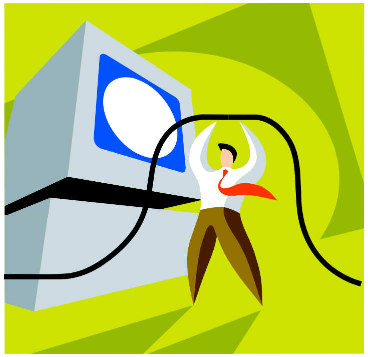
Analysis Results (4)

- Of the final 93 findings, 37 require more review because of the complexity of the code in where the potential defect exists, 5 are considered urgent (will cause data corruption or processor reset) and 37 are considered routine (worth correcting as time is available)

CSCI	CSC	All	U	R	N	TP(D)	TP(P)
CSCI B	CSC A	8	0	4	1	0.500	0.625
CSCI B	CSC B	3	0	1	0	0.333	0.333
CSCI B	CSC C	12	1	2	2	0.250	0.417
CSCI B	CSC D	18	0	4	3	0.222	0.389
CSCI B	CSC E	38	0	8	0	0.211	0.211
CSCI B	CSC F	145	3	14	29	0.117	0.317
CSCI B	CSC G	2	0	0	0	0.000	0.000
CSCI B	CSC H	14	0	5	0	0.357	0.357
CSCI B	CSC I	5	0	1	0	0.200	0.200
CSCI B	CSC J	10	0	3	0	0.300	0.300
CSCI B	CSC K	46	0	2	0	0.043	0.043
CSCI A	CSC A	29	1	7	2	0.276	0.345
Total		330	5	51	37	0.170	0.282

Findings Comparisons (1)

Data has just recently started being collected on the amount of time required to investigate DRs and provide a complete fix, so no trends can yet be provided regarding time and dollars expended on defects



Findings Comparisons (2)

- The average data loss per software discrepancy is from 2003 to 2007 was 0.331 hours
- The average data loss per software discrepancy over the past eight months is 0.579 which may be attributable to additional users of CSCI A in addition to fielding CSCI B
- A conservative estimate of savings of data loss by correcting the 5 urgent findings is 2.895 hours of data loss, which is 11.5% of the estimated annual software data loss for 2008

Findings Comparisons (3)

- Jet Propulsion Laboratory (JPL) performed a case study that found 1.2 fielded defects / LKSLOC in ground software systems they developed[1]
- Capers Jones found that CMMI Level 5 organizations deliver 1.05 defects / KLOC[2]
- Of the 12 CSCs analyzed in this study, all fall below both these defect rates except two with respect to the true positive findings

Credits: 1 – Spagnuolo, J. and Powell ,J. Defect Measurement and Analysis of JPL Ground Software: A Case Study. 2004.
2 – Jones, C. Software Assessments, Benchmarks, and Best Practices. 2000.

Findings Comparisons (4)

- The Coverity project to use static analysis technology on Open Source software produced an average of 0.605 findings / KSLOC over 96 projects using a limited set of the most easy to understand defects
- CodeSonar produced an average of 0.730 findings that were confirmed to be valid, which implies that improvements should be made in the overall quality of CSCI A and CSCI B

Acceptance Measures (1)

Measures are meant to aid in determining the impact of the technology on the project both in the expenditure of effort and the technical results of the technology

Time Spent on Tool	
One Time	Staff Hours
Training	120
Installation	4
Recurring	Staff Hours
Run Time	0.5
Result Review	2
KSLOC Analyzed	120

Finding Classification and Prioritization Counts			
	Urgent	Routine	*All Prioritizations*
True Positive	1	2	*3*
False Positive	1	0	*1*
All Classifications	*2*	*2*	*4*

Average data loss per defect

Average time to correct defect

Acceptance Measures (2)

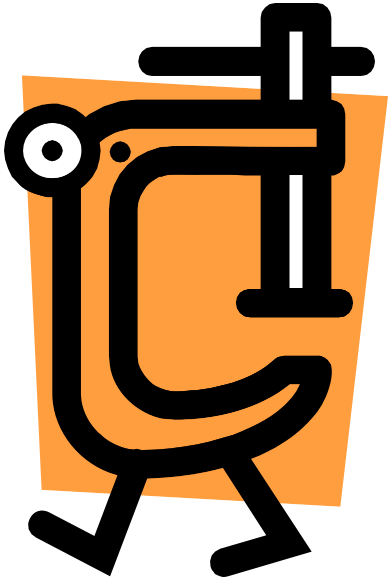
- Time will be added to sustaining engineers’ lifecycle to run CodeSonar and disposition the findings
- After relatively small initial investment for the initial tool run (compared to anecdotal comparisons to discrepancy investigation time), an estimated 10 minutes per KSLOC developed of extra time is anticipated based on preliminary use of the tool (without formal training)

Acceptance Measures (3)

- The “non-interesting” finding rate of 70% is a large number but filtering, search, and detailed contextual features of CodeSonar reduce the time per finding
- Integration of tool into the build process may also provide further savings by preventing developers from having to configure and operate the tool separately

Conclusions (1)

- Preliminary results show the tool to be easy to use and incorporate into engineering processes
- Preliminary findings give significant potential improvements in proficiency and availability on the part of software



Conclusions (2)

- As time-to-fix data becomes available a better cost trade can be made on person hours saved versus tool cost
- Selective factors may be necessary to determine where best to apply CodeSonar to balance cost and benefits